Science of the Total Environment 537 (2015) 225–234

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Getting water right: A case study in water yield modelling based on precipitation data

Natalia Pessacg [a,*], Silvia Flaherty [a], Laura Brandizi [a], Silvina Solman [b], Miguel Pascual [a]

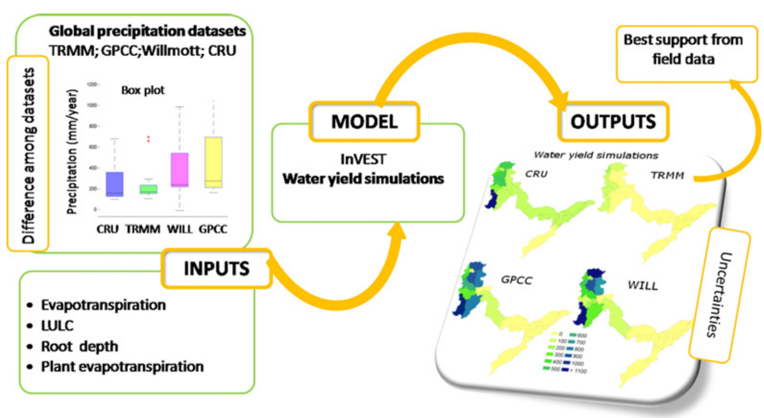
[a] Centro Nacional Patagónico (CENPAT/CONICET), Blvrd. Brown 2825, U9120ACF Puerto Madryn, Chubut, Argentina

[b] Centro de Investigaciones del Mar y la Atmósfera (CIMA/CONICET-UBA), DCAO/FCEN, UMI IFAECI/CNRS, Ciudad Universitaria Pabellón II Piso 2, C1428EGA Buenos Aires, Argentina

HIGHLIGHTS

- Differences in precipitation over the Andes in Central Patagonia are striking.
- The sensitivity of water yield modelling to precipitation inputs is large.
- Small differences among datasets translate into large differences in water yield.
- It is essential to evaluate the uncertainty related to input when modelling ES.
- Using the best precipitation data is fundamental for modelling hydrological ES.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 11 March 2015
Received in revised form 29 July 2015
Accepted 29 July 2015
Available online 15 August 2015

Editor: D. Barcelo

Keywords:
Precipitation data
Ecosystem services modelling
Water yield
Uncertainties
Chubut River Basin

ABSTRACT

Water yield is a key ecosystem service in river basins and especially in dry regions around the World. In this study we carry out a modelling analysis of water yields in the Chubut River basin, located in one of the driest districts of Patagonia, Argentina. We focus on the uncertainty around precipitation data, a driver of paramount importance for water yield. The objectives of this study are to: i) explore the spatial and numeric differences among six widely used global precipitation datasets for this region, ii) test them against data from independent ground stations, and iii) explore the effects of precipitation data uncertainty on simulations of water yield. The simulations were performed using the ecosystem services model InVEST (Integrated Valuation of Ecosystem Services and Tradeoffs) with each of the six different precipitation datasets as input. Our results show marked differences among datasets for the Chubut watershed region, both in the magnitude of precipitations and their spatial arrangement. Five of the precipitation databases overestimate the precipitation over the basin by 50% or more, particularly over the more humid western range. Meanwhile, the remaining dataset (Tropical Rainfall Measuring Mission — TRMM), based on satellite measurements, adjusts well to the observed rainfall in different stations throughout the watershed and provides a better representation of the precipitation gradient characteristic of

Abbreviations: InVEST, Integrated Valuation of Ecosystem Services and Tradeoffs; ES, ecosystem services; CHB, Chubut River Basin; UCH, Upper Chubut River Basin; MCH, Middle Chubut River Basin; LCH, Lower Chubut River Basin; CRU, Climate Research Unit; TRMM, Tropical Rainfall Measuring Mission; GPCC, Global Precipitation Climatology Centre; TRMMv6, Tropical Rainfall Measuring Mission version 6; TRMMv7, Tropical Rainfall Measuring Mission version 7; INTA, Agriculture Technology National Institute; LULC, land use/land cover; FAO, Food and Agriculture Organization of the United Nations.

* Corresponding author.
E-mail addresses: pessacg@cenpat-conicet.gob.ar (N. Pessacg), flaherty@cenpat-conicet.gob.ar (S. Flaherty), brandizi@cenpat-conicet.gob.ar (L. Brandizi), solman@cima.fcen.uba.ar (S. Solman), pascual@cenpat-conicet.gob.ar (M. Pascual).

http://dx.doi.org/10.1016/j.scitotenv.2015.07.148

the rain shadow of the Andes. The observed differences among datasets in the representation of the rainfall gradient translate into large differences in water yield simulations. Errors in precipitation of +30% (−30%) amplify to water yield errors ranging from 50 to 150% (−45 to −60%) in some sub-basins. These results highlight the importance of assessing uncertainties in main input data when quantifying and mapping ecosystem services with biophysical models and cautions about the undisputed use of global environmental datasets.

1. Introduction

Freshwater flows are essential for ecosystems, agriculture, industry, human consumption, hydropower, fisheries, and recreation. Not surprisingly then, recent developments in ecosystem services modelling for resource management have focused on water yield as a key ecosystem input.

Patagonia is rich in water resources, though population growth, and urban and industrial development are taking a toll on the capacity of watersheds to deliver water related benefits to people. Besides that, climate change scenarios predict a decrease in annual precipitations and an increase in temperature for most of the region (Nuñez et al., 2008; Garreaud and Falvey, 2009), making the situation worse. For instance, freshwater is becoming a limited resource for different urban areas in the region, especially along the arid Atlantic coast where most people live.

Here we conduct a modelling analysis of water yields for the Chubut River basin in Patagonia, Argentina. The Chubut River is ideal for such an analysis because it is a relatively small river basin, and is yet important to as many as 50% of the population in the Chubut Province. This river, as most of the rivers that flow across the arid Patagonia, is fundamental for agricultural irrigation and water supply for human consumption. This basin is also an important case study in terms of ecosystem services (ES) due to the existence of a number of regional projects aimed at creating hydroelectric dams which will provide irrigation water for new agricultural areas (Plan Director, 2013).

All ecosystem services models involving freshwater derive in part from precipitation data as a critical input. Precipitation is a key driver of the hydrological cycle, as well as one of the most difficult variables to measure because of its high spatial heterogeneity and temporal variability (Junzhi et al., 2012). Accurate measurements of precipitation are important to environmental scientists, as well as to a wide range of decision makers related to faming, industry, emergency management, urban areas, and natural areas (Ebert et al., 2007). However, estimating precipitation is difficult in many parts of the world due to the logistic complexity and high costs of establishing and maintaining the required infrastructure (Yilmaz et al., 2005). Patagonia is not an exception and the spatial characterization of precipitation based on actual field data is sparse at best, due to nonexistent ground-based radars and a poor network of rain gauges. When precipitation data at the scale of a watershed are required the only alternative is to use global precipitation datasets. Such data sets, available in different resolutions, are based on ground observations, satellite estimation, a combinations of both, or outputs from general circulation models. Whereas many studies have found that global precipitation datasets provide a good representation of temporal trends and global-scale spatial distribution, they often exhibit marked differences among themselves at the regional scale (Getirana et al., 2011). Getirana et al. (2011) compared six daily and sub-daily precipitation datasets and their impacts on the water balance of the Negro River Basin (Amazonas). They found that gauge-based data are the most accurate; however some satellite and model-based datasets can reproduce the water cycle at the basin scale and monthly time step fairly well. In addition to this, Hamel and Guswa (2015) found that the uncertainties introduced by errors in climate input data are significant and spatially heterogeneous, affecting the spatial distribution within the Cape Fear watershed of areas with high and low water yields.

Whereas the above studies suggest that precipitation data can be highly uncertain, systematic studies of the consequences of this uncertainty for resource management priorities are rare. In Patagonia, the consequences of this lack of analysis might be worse due to data scarcity and the vastness of the region.

In this context, the aim of this study was to analyze the agreement among different global precipitation datasets and with data from available independent ground stations and to explore the effect of precipitation data uncertainty on simulations of water yield over a Patagonian pilot watershed, the Chubut River Basin (CHB). Because so much interest in water yield stems from assessments of ES, we used the model InVEST to generate spatially explicit estimates of water yield in different sub-basins of the CHB.

The paper is organized as follows: Section 2 describes i) the InVEST model used for simulating water yield, ii) the precipitation datasets available for the CHB, and iii) the characteristics of the CHB. In Section 3, we present a comparison of the precipitation datasets and the sensitivity of water yield in the Chubut watershed to the uncertainties in precipitation data using the InVEST model. We also present a calibration of the model to the ecohydrological parameter Z. Finally, Section 4 includes discussion and conclusions.

2. Methodology

2.1. The Chubut River Basin as a case study

The Chubut River originates in the western extra Andean region of Patagonia (Rio Negro province) and flows for about 800 km, first south and then east across the Chubut province and into the Atlantic Ocean (Fig. 1). The basin reaches an altitude of 2300 m in its source, Nacimiento, at an altitude of 950 m, is the higher and westernmost flow gauge station used for this study. The Chubut River's hydrograph is characterized by two peak discharges, one in spring from snowmelt and the other in the fall from rainfall.

The CHB basin has a total area of 57,400 km^2 and is divided in 3 major sub-basins, upper (UCH), middle (MCH) and lower (LCH), themselves divided in a total of 24 sub-basins (Fig. 1). The lower basin is the most populated, concentrating more than 50% of the Chubut province population. The Chubut River is the only water supply for over 200,000 people (Commendatore and Esteves, 2004).

One multipurpose dam, Florentino Ameghino, with a capacity of 1500 hm^3, was built in 1968 for energy production and to regulate flow of the lower basin for flood control, irrigation, and water provision (Plan Director, 2013).

2.2. Precipitation data

A number of global or quasi-global precipitation datasets, with different spatial and temporal resolutions, and generated using different methodologies are available. For the purpose of this study and considering the dimensions of the basin, only datasets with the higher spatial resolution were selected. In total, six mean annual precipitation datasets were analyzed (Table 1): the Climate Research Unit (CRU); the Willmott from the University of Delawere; the Global Precipitation Climatology Centre (GPCC), and three different versions of the precipitation datasets produced by the Tropical Rainfall Measuring Mission (TRMM), which is

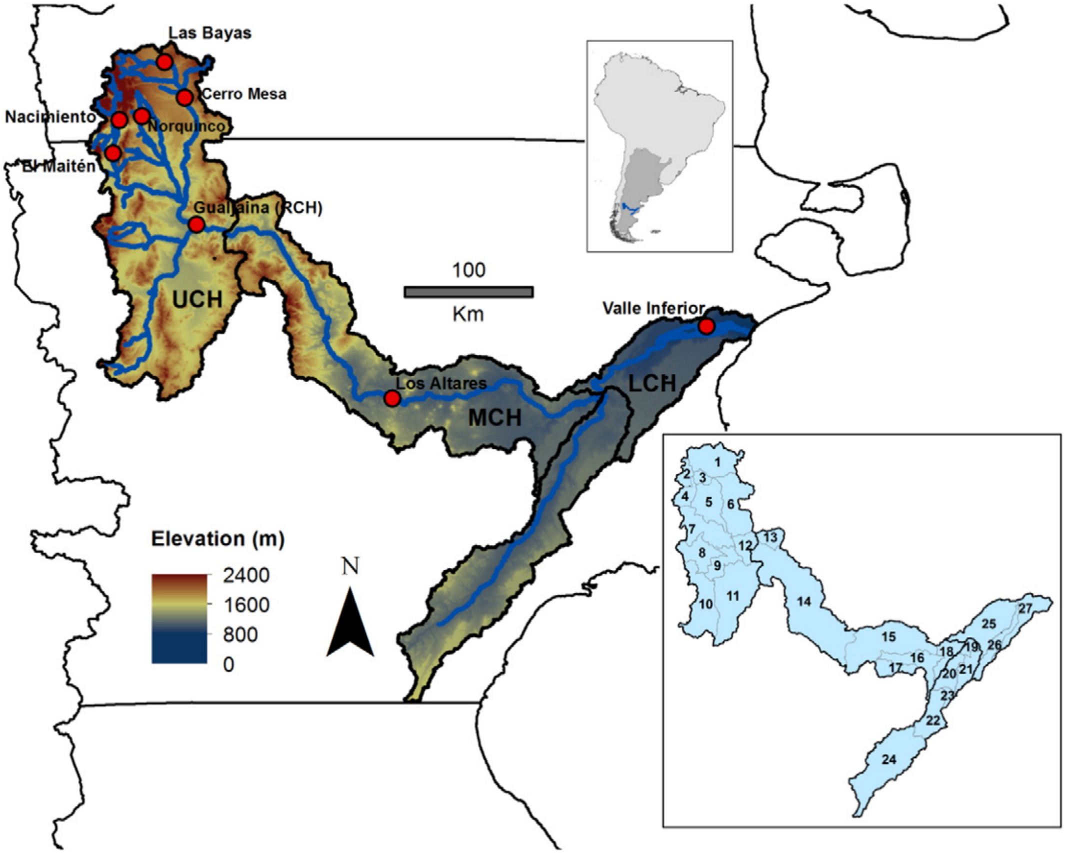


Fig. 1. Location and topography (metres above sea level) of the Chubut river basin and the three major sub-basins: upper Chubut (UCH), middle Chubut (MCH) and lower Chubut (LCH). The red points show the location of the rain gauge stations. The numbers in the lower-right panel indicate the location of the 24 sub-basins the watershed is divided in.

a joint mission between NASA and the Japan Aerospace Exploration Agency.

The first three datasets are generated using ground data (i.e. rain gauge stations) and different interpolation methods. The spatial resolution of these datasets is 0.5° × 0.5° (approximately 50 km). The three TRMM versions are based on satellite measurements and differ on the algorithms used. The first version used in this work – TRMM – uses the algorithm 3B42, which generates, calibrates and combines the microwave and infrared precipitation estimations with a 3 hour temporal resolution. This product is then rescaled to monthly data (Huffman et al., 2007). The other two versions, TRMMv6 and TRMMv7, use the algorithm 3B43, which combines the estimates of the 3B42 algorithm with the global rain gauge data gathered by the Global Precipitation Climatology Centre (GPCC) (Huffman, 2012). TRMMv7 changes the geolocation to latitude and longitude, adds many variables, layers and gauge relative weighting in regard with TRMMv6. The three satellite precipitation datasets have a 0.25° × 0.25° spatial resolution. The TRMM satellite orbits at an inclination of 35° and the spatial coverage extends from 50°N to 50°S latitude. We analyze the mean annual precipitation for the period 1998–2008, a time span common to all datasets, for a section of Patagonia including the CHB basin and large enough to appreciate regional precipitation patterns (42 to 50°S and 73 to 60°W).

The uncertainty in the grid observations was evaluated as the spread among the different precipitation datasets. This methodology allows determining the degree of agreement among the precipitation datasets. The spread was calculated as the standard deviation of the 11-year mean of each precipitation dataset with respect to the mean precipitation dataset ensemble.

In terms of ground stations, there was data available for eight rain gauges over the basin (Fig. 1). This ground stations are operated by the National Meteorological Service, the Agriculture Technology National Institute (INTA) and the National Secretary of Water Resources. The rain gauge stations are regularly subjected to quality control, however, they may be subject to errors due strong winds or topography (Groisman

Table 1
Precipitation datasets considered as input in the InVEST model.

Dataset	Reference	Spatial coverage	Temporal resolution	Spatial resolution	Source
CRU University of East Anglia Climate Research Unit	Mitchell and Jones (2005)	Global	1901–2009	0.5° × 0.5°	Rain gauge CRU TS3.1
GPCC Global Precipitation Climatology Centre	Schneider et al. (2011)	Global	1901–2014	0.5° × 0.5°	Rain gauge
Willmott University of Delawere	Legates and Willmott (1990)	Global	1900–2009	0.5° × 0.5°	Rain gauge
TRMM Tropical Rainfall Measuring Mission	Huffman et al. (2007)	50°N to 50°S	1998–2009	0.25° × 0.25°	Satellite (3B42 algorithm, version 6)
TRMMv6 Tropical Rainfall Measuring Mission	Huffman et al. (2007)	50°N to 50°S	1998–2010	0.25° × 0.25°	Satellite + rain gauge (3B43 algorithm, ver. 6)
TRMMv7 Tropical Rainfall Measuring Mission	Huffman et al. (2007), Huffman (2012)	50°N to 50°S	1998–2010	0.25° × 0.25°	Satellite + rain gauge (3B43 algorithm, ver. 7)

and Legates, 1994; Daly et al., 1994; Luo et al., 2005). Nevertheless, and given the lack of other sources of data, they have been considered as reference in this study.

We chose to focus on mean precipitation over a decade because ecosystem services models focus on average yields when identifying specific regions that might be targeted as valuable for different ecosystem services—in this case the delivery of water.

2.3. Annual water yield simulation

We used InVEST, a suite of models to map and valuate the goods and services from nature that sustain and fulfil human life (http://www.naturalcapitalproject.org/InVEST.html). These models are spatially explicit and were developed by the Natural Capital Project with the aim of identifying where investment may enhance human well-being and nature. We used the InVEST water yield model (Reservoir Hydropower Production) to analyze the total annual water yield in the CHB. The basic output of this module is a gridded map of water yield.

Annual water yield (Y) for each pixel on the landscape (x) is calculated as:

$$\mathrm{Yx} = \left(1 - \frac{\mathrm{AETx}}{\mathrm{Px}}\right) * \mathrm{Px}$$

where AET is the annual actual evapotranspiration and P is the annual precipitation.

The relationship between AET and P is based on the Budyko curve (Budyko, 1974) as follow:

$$\frac{\mathrm{AETx}}{\mathrm{Px}} = 1 + \frac{\mathrm{PETx}}{\mathrm{Px}} - \left[1 + \left(\frac{\mathrm{PETx}}{\mathrm{Px}}\right)^{w}\right]^{1/w}$$

where PETx is the potential evapotranspiration for pixel x.

The empirical parameter w characterizes the natural climatic-soil properties and its expression is:

$$w_x = Z * \frac{\mathrm{AWCx}}{\mathrm{Px}} + 1.25$$

where AWC is the plant available water content and Z is an empirical parameter. Potential evapotranspiration is defined as:

$$\mathrm{PETx} = \mathrm{Kc(lx)} * \mathrm{ETox}$$

where ETo is the reference evapotranspiration and Kc is the plant evapotranspiration coefficient associated with the LULC (lx) in each pixel. Details about the model can be found in Sharp et al. (2014).

The water yield model requires several inputs from different sources (see Table 2). ETo was obtained from the Food and Agriculture Organization of the United Nations (FAO).

The FAO estimates ETo using the Penman–Monteith method (Allen et al., 1998). The LULC, root restricting layer depth, plant available water content, and soil depth maps were all provided by INTA (GeoInta, http://geointa.inta.gov.ar). The watershed and the subwatersheds polygons were calculated using Arc Hydro tools (ArcGIS 10.1) and checked against existing maps from the National Secretary of Water Resources of Argentina (Giraut et al., 2010).

We first explored annual water yield in different sub-basins using each of the six mean annual precipitation gridded datasets. We analyzed the overall geographic pattern of water yield over the basin, and compared the specific regional patterns produced by different precipitation datasets.

A validation analysis with streamflow data was performed for one station in the basin, Los Altares, located in the MCH (Fig. 1). Due to the lack of water yield measurement in the basin, streamflow data at Los Altares was compared with the aggregated annual water yield predicted from all upstream sub-basins contributing to the point in the river where the gauge is.

To further analyze the sensitivity of water yield from InVEST to precipitation uncertainty, another round of simulations was performed using controlled precipitation errors around observed data from the output of the TRMM algorithm 3B42 (TRMM in Table 1). For a group of seven sub-basins (those with water yield values higher than 10^6 m^3 yr^{-1}) and for a group of four pixel types with contrasting conditions of LULC, mean annual precipitation was systematically varied around observed values and the response in annual water yield was recorded. We used the TRMM annual mean precipitation dataset plus systematic errors of $\pm$ 10%, $\pm$ 20%, and $\pm$ 30% as inputs to InVEST. The result of this analysis is a plot of relative change in water yield for different systematic errors in precipitation (Paturel et al., 1995; Goyal, 2004; Xu et al., 2006 among others).

2.3.1. Ecohydrological parameter Z

The ecohydrological parameter Z is by definition an empirical parameter which captures the climate seasonality, rainfall intensity, and topography characteristics in the basin. However, the parameter Z can be also used as a calibration constant to correct the simulations for the effects of specific processes the InVEST model does not consider, such as the frequency of annual events, the sub-parcel spatial variability of soil water storage capacity, and the synchronicity of the energy-precipitation cycles (Milly 1994; Kareiva et al., 2011).

Three methods have been proposed to calculate Z (Sharp et al., 2014): i) as a linear function of the number of precipitation events per year (Donohue et al., 2012); ii) using global estimations of w (Xu et al. 2013; Liang and Liu 2014), iii) estimation from calibration using streamflow data.

For the baseline run, we used a Z value of 5 which was calculated as the mean value of one-fifth of the number of rain days per year in each rain gauge (following Donohue et al., 2012). Alternatively, a calibration analysis was performed with respect to the parameter Z using the

Table 2
InVEST data needs: Input files for the water yield module (InVEST 3.0.1).

Input	Description/units	Format	Source
Root restricting layer depth	Soil depth at which root penetration is strongly inhibited because of physical or chemical characteristics/mm	GIS raster	INTA soil map http://geointa.inta.gov.ar
Precipitation	Average annual precipitation/mm yr^{-1}	GIS raster	Precipitation datasets (see Table 1)
Plant available water content	Fraction of water that can be stored in the soil profile that is available for plants' use/mm	GIS raster	INTA & Ministry of Agriculture, Food & Fisheries, British Columbia (Canada)
Reference evapotranspiration	Average annual reference evapotranspiration/mm	GIS raster	FAO (http://www.fao.org)
Land use/land cover (LULC)	Land use/land cover map	GIS raster	INTA http://geointa.inta.gov.ar
Watersheds	Main watershed	GIS polygon	GIS toolsArcHydro
Subwatersheds	Subwatersheds within the main watersheds	GIS polygon	GIS tools/ArcHydo
RootDepth	Maximum root depth for vegetated LULC classes/mm	Integer number	Canadell et al. (1996)
Plant evapotranspiration coefficient	Plant evapotranspiration coefficient for each LULC class	Decimal number between 0 and 1.5	FAO (http://www.fao.org)
Seasonality factor (Z)	Seasonal distribution of precipitation	Integer number between 1 and 30	Donohue et al. (2012)

difference between annual streamflow observed at Los Altares (MCH, Fig. 1) and estimated annual water yield upstream of that point as the calibration criteria. For each of the precipitation datasets, simulations using alternative values of Z within the full range of possible values (1–30) were performed. We report the value of the calibration criteria, as the percent difference between annual streamflow and water yield, as a function of Z values for different precipitation datasets.

3. Results

3.1. Mean annual precipitation

All six different datasets capture the west–east gradient over Patagonia (Fig. 2) of the annual mean precipitation. However, they exhibit large differences in the location and extent of precipitation maxima over the Andes (southward 40°S). These differences are associated with the intrinsic difficulties in representing the strong precipitation gradient over this area, where annual mean precipitation ranges from 6000 mm yr^{-1} on the Chilean coast to less than 300 mm yr^{-1} in Argentina, just a few tens of kilometres downstream from the Andes (Smith and Evans, 2007).

Over the CHB, the precipitation datasets show maximum values ranging from around 1000 mm yr^{-1} over the west side of the basin to less than 200 mm yr^{-1} over the mid and lower basins. The largest differences among datasets are located over the west border of the basin, in consonance with the aforementioned problems to represent maxima over the Andes.

The main characteristics of each dataset over the CHB are summarized in a Fig. 3. The annual mean precipitation for the 11-year period over the CHB ranges from around 150 mm yr^{-1} for CRU and TRMM (algorithm 3B42) to 450 mm yr^{-1} for TRMMv6 (algorithm 3B42). This variability in the mean value is larger than the mean precipitation in the LCH and MCH sub-basins.

Moreover, there are large differences in the representation of the 25th and 75th percentile. TRMM and CRU show the smallest spread among percentiles associated with a smaller west–east precipitation gradient.

Finally, the spread among precipitation datasets over the entire basin was calculated (Fig. 4). In general, large spread values were observed over the entire basin, reaching 300 mm yr^{-1} in some UCH sub-basins, which represent 50% of the mean annual precipitation.

Summarizing, there is a large spread among precipitation datasets in terms of annual precipitation averages over the 11-year period. The largest differences are registered over the western side of the basin, which extends over the strong precipitation gradient characteristic of the rain shadow east of the Andes. In particular, these differences are associated with how far east high Andean precipitations are projected over the basin by different datasets (Fig. 2).

The different datasets have very different performances when compared to data from rain gauge stations (Fig. 5). This kind of analysis are best done by first interpolating the point-based rain gauge data to areal precipitation (at the same scale of the precipitation datasets) using interpolations techniques (i.e. Thiessen polygon or Kriging). However, the paucity of rain gauge data for the CHB did not allow us to properly estimate rainfall on an areal basis. We therefore choose a straight comparison of point rain gauge values with the nearest point corresponding to the precipitation dataset (Fig. 5). This method can introduce local errors due to the low spatial resolution of the gridded precipitation datasets. The insufficient and unevenly distributed rain gauge network can also introduce errors in the comparison with precipitation datasets due to the effect of wind and topography. The effect of strong winds in precipitation, in particular during snow events, is

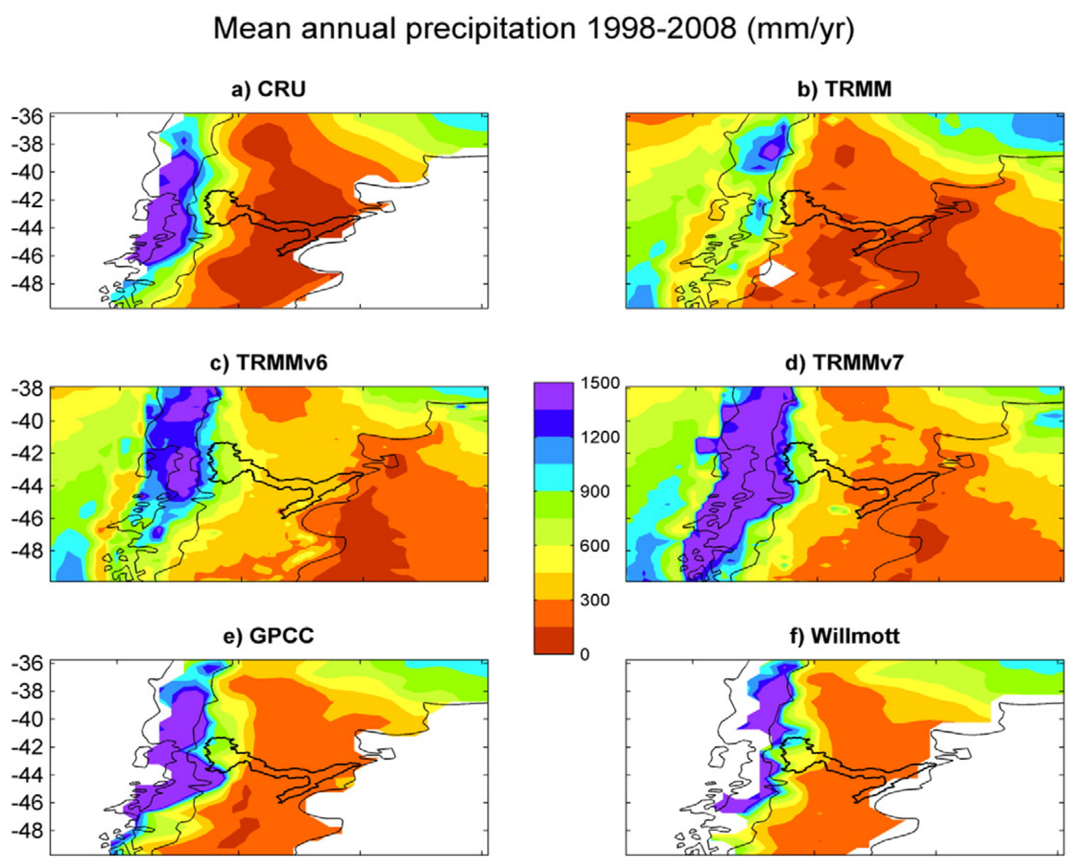


Fig. 2. Mean annual precipitation 1998–2008 (mm $year^{-1}$) for six dataset (a) CRU, b) TRMM, c) TRMMv6, d) TRMMv7, e) GPCC, f) Willmott. The Chubut River Basin bounds are depicted in the maps.

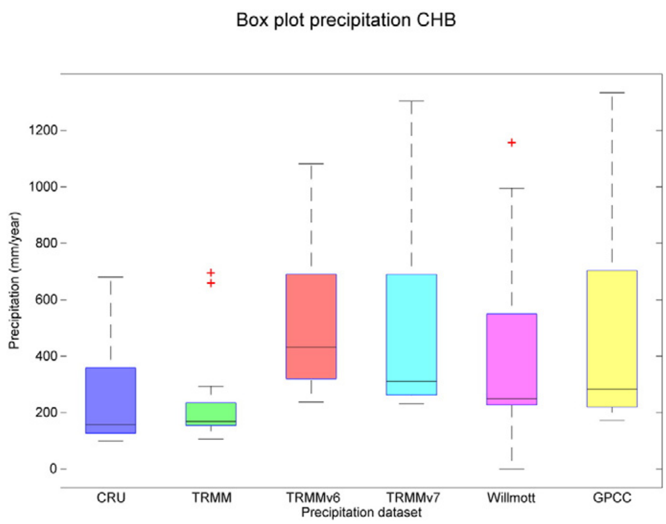


Fig. 3. Box plots for the mean annual precipitation in the Chubut river basin for the six observational dataset.

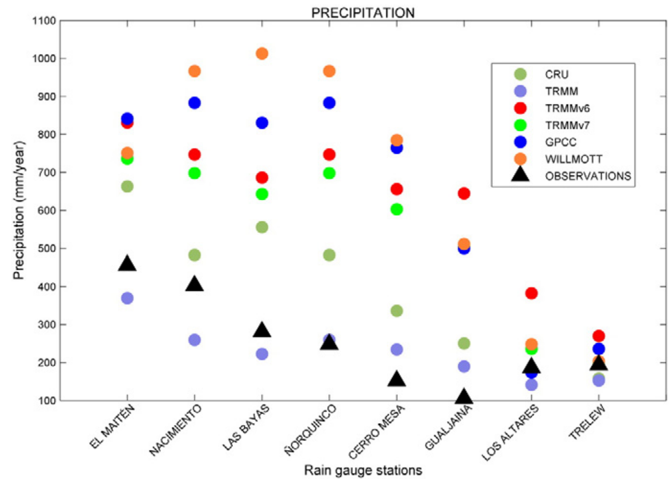


Fig. 5. Precipitation (averages for 1998–2008, in mm yr^{-1}) for the different datasets in regards to the rain gauge stations.

generally misrepresented (Groisman and Legates, 1994). Furthermore, since rainfall increases sharply with altitude, precipitation over areas with complex topography tends to have systematic errors (Daly et al., 1994; Luo et al., 2005). Despite these possible sources of error, gridded precipitation datasets provide an overall picture of biases and differences among databases and their general effects as inputs to ES models.

In general, the datasets overestimate the precipitation with respect to rain gauge data in the more humid stations, particularly over the west of the CHB near the Andes. Willmott and GPCC present the largest overestimations for most of the stations over the UCH sub-basin. In general, the datasets show small biases in the MCH and LCH sub-basins, represented by the rain gauge stations "Los Altares" and "Trelew", respectively.

TRMM (with the 3B42 algorithm) is the database that best adjusts to the observed rainfall in different stations throughout the watershed.

Interestingly, large differences in precipitation are portrayed by different TRMM versions. Both versions of TRMM that use the 3B43 algorithm (TRMMv6 and TRMMv7), which incorporate rain gauge data, overestimate the precipitation much more than the TRMM, which used the 3B42 algorithm, without rain gauge data. This suggests that in the CHB, where rain gauge data are scarce, the result of incorporating

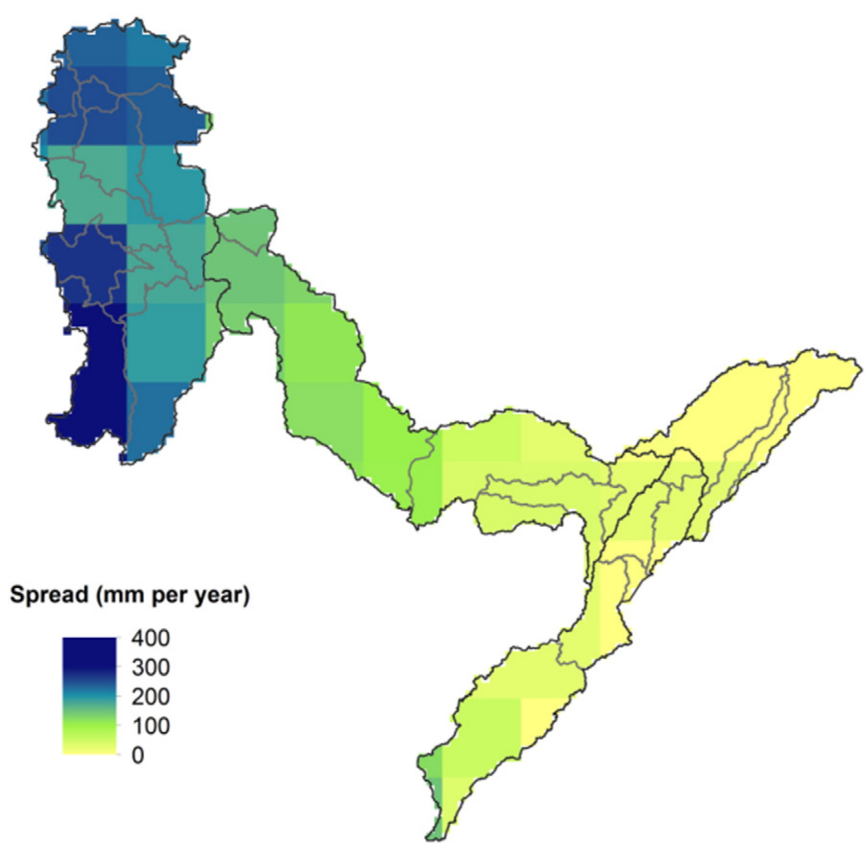


Fig. 4. Spread among precipitation datasets with respect to the mean precipitation dataset ensemble. (Unit: mm yr^{-1}).

these data into the satellite estimations of precipitation may decrease the quality of the dataset.

3.2. Sensitivity of water yield to precipitation errors

Because precipitation is a major driver of water yield, it comes as no surprise that different precipitation datasets produce large differences in simulated water yield (Fig. 6). These simulations where performed using the same inputs and basin configuration; but precipitation data from each of the datasets listed in Table 1.

All datasets agree in allocating maximum water yields to the upper basin, close to the Andes, but with very important differences not only in the magnitude but in the location of water yield. In agreement with the largest differences in the representation of precipitation, the largest water yield differences are observed among sub-basins of the UCH. In particular for the sub-basins of the UCH with higher water production the water yield simulated varied around 40–50% depending on the precipitation data used as input.

Simulations performed using precipitation from CRU and TRMM, the ones that best reproduce the rain gauge data (Fig. 5), show the lowest values of water yield over the UCH. These two same data sets, however, produce a very different geographic pattern of water yield at the sub-basin level (Fig. 7), particularly for the most western sub-basins in the UCH (4, 5, 7, and 10). In fact, the relative ranking of sub-basins in

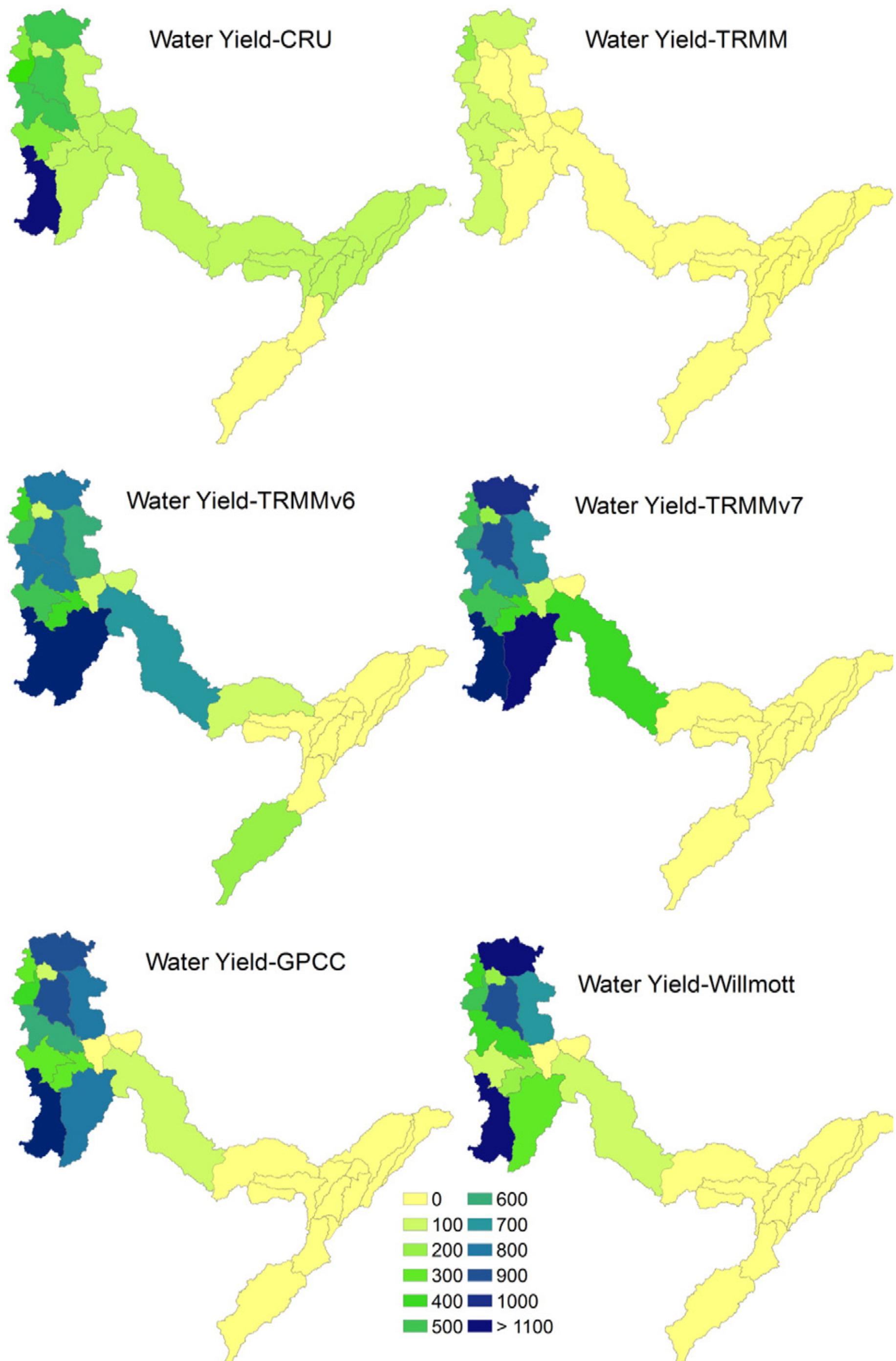


Fig. 6. Sensitivity of water yield (10^6 m^3 yr^{-1}) modelling with InVEST to different precipitation datasets.

terms of water yield is widely discordant depending on which precipitation data were used.

On the other hand, simulations forced with GPCC, Willmott, and both TRMMv6 and TRMMv7 which used the 3B43 algorithm, double the water yield magnitude in upstream sub-basins when compared with TRMM (3B42 algorithm). The differences in water yield estimation over the middle and lower basin, on the other hand, are not as large.

The validation for Los Altares (Fig. 8), located in the MCH basin (Fig. 1), shows that the water yield simulated using the TRMM precipitation dataset and the InVEST model presents the lowest bias in regard to the streamflow data for this location (23%), followed by the simulation using CRU, with a bias of 78%. The rest of simulations using the other precipitation datasets overestimate water yield by more than 300%. These results suggest that the validation using streamflow data and global precipitation datasets are consistent throughout the basin.

The specific sensitivity of water yield to systematic errors in precipitation varies both among representative sub-basins and pixels with different LULC class selected throughout the watershed (Fig. 9). The sensitivity curves show a strong dependence of water yield on changes in precipitation. Errors of +30% in precipitation lead to changes in water yield ranging from 50 to 150%, depending on the location of the sub-basin. Errors of −30% in precipitation led to changes between −45 to −60%, in the sub-basins with the highest values of water yield (Fig. 9a). The differences between the impact of negative and positive errors are inherent to the shape of the Budyko curve, used by InVEST to model the relationship between the annual actual evapotranspiration and precipitation. The sensitivity to positive precipitation errors depends to a great extent on the LULC (Fig. 9b). The sensitivity to precipitation errors is larger for LULC categories with higher values of Kc, a parameter related to the potential evapotranspiration of different LULC classes.

The spread among annual mean precipitation from different datasets (Fig. 3) is in the same order of magnitude as the idealized errors introduced in these simulations. That means that the plausible errors of ±30% that a model user can introduce when an arbitrary rainfall dataset is considered may lead to errors as large as −45 to 150% in the water yield estimation at the sub-basin scale, highlighting the importance of seriously scrutinizing precipitation and other input parameters by comparison with field data and sensitivity analyses.

The Tecka sub-basin (number 10 in Fig. 1) provides an interesting example of critical data voids. With the exception of TRMM (algorithm

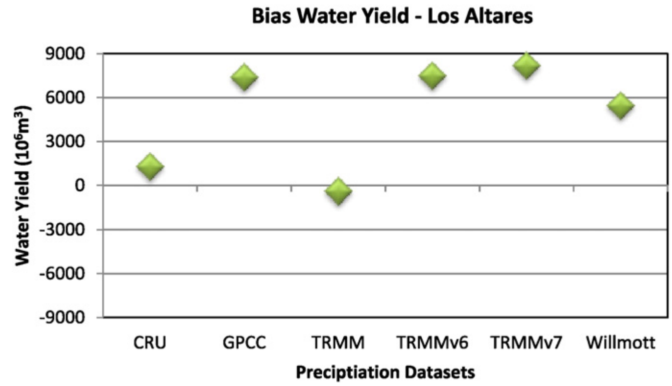


Fig. 8. Bias between annual water yield modelling and streamflow data in Los Altares (for the period 1998–2008, in $10^6 m^3 yr^{-1}$) for the different datasets. To calculate the predicted annual water yield for Los Altares we added up the annual water yield for the sub-basins upstream the gauge stations.

3B42), different precipitation datasets indicate this is the area where most water is produced in the basin (Fig. 6), and also the area with the largest spread among precipitation datasets (Figs. 2 and 3). In addition, it is the sub-basin presenting the largest sensitivity to precipitation errors (Fig. 9a). Meanwhile, no rain gauge stations are available to validate the actual precipitation in this sub-basin.

3.3. Calibration to the ecohydrological parameter Z

Besides providing a calibration of the model to parameter Z, the analysis also provides a detailed analysis of the sensitivity of water yield to Z (Fig. 10). In general, changes in Z have a large impact on water yield simulations. Extreme values of Z (1 and 30) lead to differences in water yield simulations of 148% and 47% respectively when compared with the baseline run (Z = 5, TRMM precipitation, Fig. 10). Lower Z values have larger effects on water yield simulations than higher values. This is particularly important in arid basins such as the CHB, where the number of rain events – and therefore the Z value – is low.

In terms of calibration, a full agreement between water yield simulations and streamflow data is found for Z = 3 when TRMM precipitation data were used (Fig. 10). This Z value, which is comparable to that calculated considering the number of rain events (Z = 5), produces a reduction in estimated water yield of 37% with respect to the baseline run.

On the other hand, water yield for the CRU precipitation data, which provided a poor fit to streamflow data for the baseline value Z = 5, provides a full agreement for Z = 15.

Meanwhile, no value of Z within the range was found that allows for a calibration of the model for the remaining precipitation datasets (i.e. GPCC, TRMMv6, TRMMv7, Willmott). For example, a value of Z = 30, the maximum value that the parameter can take, leads to biases in Los Altares larger than 165%.

4. Discussion and conclusions

The sensitivity of annual water yield modelling in the Chubut River Basin to different precipitation inputs was striking, with errors in precipitation of +30% (−30%) leading to water yield errors ranging from 50 to 150% (−45 to −60%) in some sub-basins.

While these results are particular to Chubut River basin they have been found in other basins that span steep gradients in elevation and precipitation (Sánchez-Canales et al., 2012). In CHB these conditions are met because the headwaters of the Chubut River are located over the strong rainfall gradient at the rain shadow of the Andes. The Andes strongly affect the regional climate by blocking the disturbances embedded in the westerly flow, producing the precipitation over this area and influencing regional wind patterns and precipitation (Insel

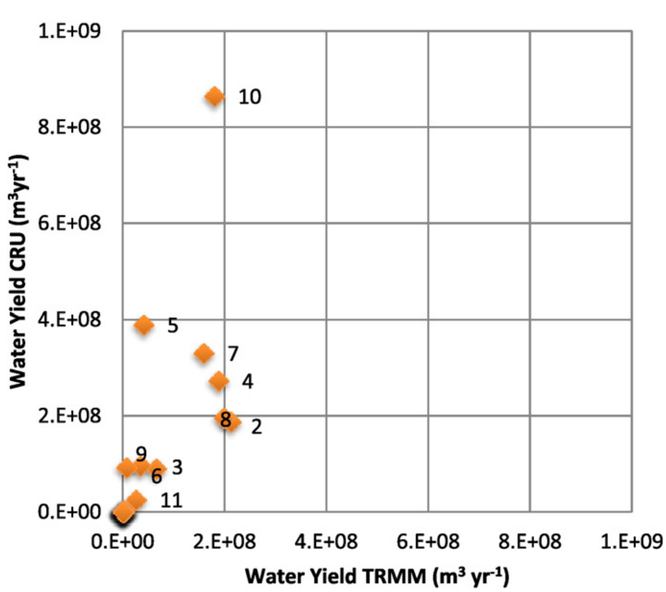


Fig. 7. Scatter Plot of water yield by sub-basin for CRU and TRMM (3B42 algorithm) precipitation datasets. Points and labels represent different sub-basin in the watershed (see Fig. 1).

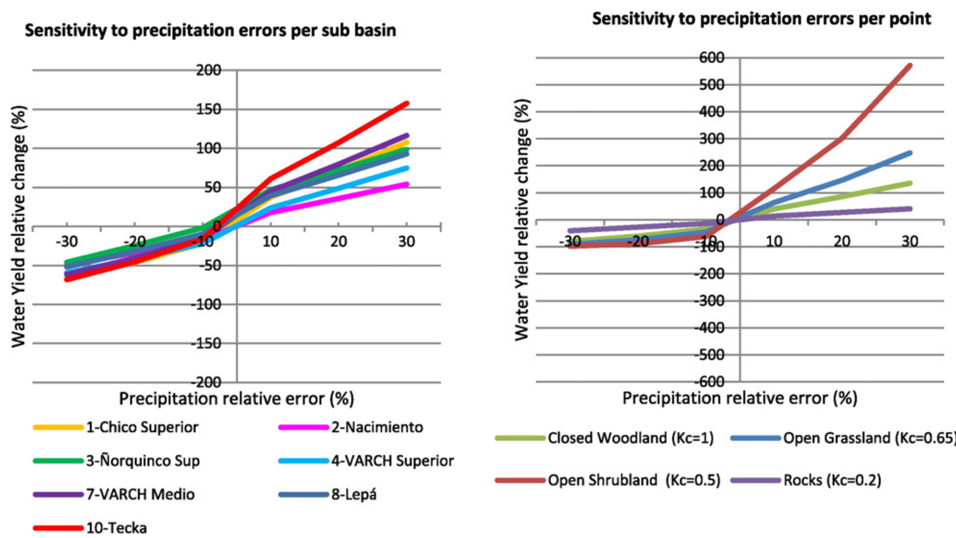


Fig. 9. Sensitivity to precipitation errors, (a) per sub-basin and, (b) per point related with different land use categories (in the legend, in brackets, is the Kc value for each land use category).

et al., 2009). Uplift on the west side of the Andes leads to hyper humid conditions, while down slope subsidence dries the eastern plains leading to arid and highly evaporative conditions (Garreaud et al., 2013). Some databases provide a better representation of this precipitation gradient, but most of them extend the maximum rainfall values to the east and, consequently, overestimate the precipitation over the west side of the CHB. The spread among precipitation datasets is larger than 30% over this part of the basin. The comparison of precipitation datasets with rain gauge data over the CHB indicates that the TRMM (algorithm 3B42) dataset is the most accurate. However, it is important to remark that the distribution of precipitation over this area is poorly known, especially over the rainfall maximum area, because of the paucity of rain-gauge stations in this inaccessible region (Garreaud et al., 2013).

In the end, small differences among datasets in the extent of precipitation maxima over the Andes translate into large differences in the estimation of the amount of water received by the Chubut watershed. Moreover, small differences among the same datasets in the spatial distribution of precipitation maxima over the Andes produce very different estimates of the relative amount of water received by different watersheds and of their relative water yield.

As was evident from the results for the Tecka sub-basin, where precipitation varied around 40% among datasets and water yield around 60% depending on the precipitation dataset used as input, it is essential when producing ecosystem service production functions to understand the different sources of uncertainty.

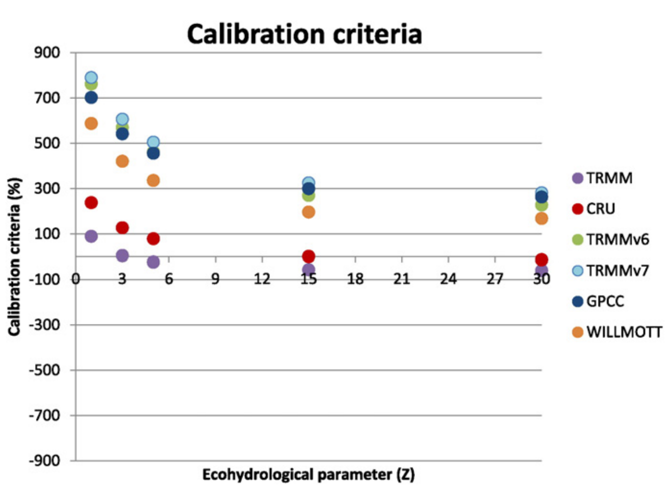


Fig. 10. Calibration criteria for the period 1998–2008, as a function of Z value for the different precipitation datasets.

The results of this study also suggest that water yield predictions are highly sensitive to the natural variation in land use land cover characteristic of this watershed in particular and Patagonian watersheds in general, which typically include some kind of gradient including forests, grasslands, and shrublands. In particular, a strong water yield dependence on the evapotranspiration coefficients characteristic of different vegetation types was evident. Further research is needed to fully explore the effects of changes in LULC not only on water yield predictions but also on other ES which might have a stronger dependence on these changes (i.e. sediments retention, nutrient transport).

The analysis of the ecohydrological parameter Z shows that in this kind of arid basins, small differences in the selection of Z lead to large differences in water yield simulations. This is due to the model being more sensitive to variations at lower Z values. Previous researchers have found that for other basins with larger number of annual precipitations events – and therefore larger Z values – the sensitivity of the model to Z is not relevant (Hamel and Guswa, 2015; Sánchez-Canales et al., 2012). Our specific results show that the effect of large precipitation errors on model fit cannot be absorbed by the parameter Z. Adjusting the value of Z allows for a discrimination of the effects of errors due to model structure, parameter estimation, and climate variables (Hamel and Guswa, 2015).

Considered together, the results highlight the importance of not using global datasets on precipitation and other inputs to ES models uncritically, for errors in mapping ES could be large, both in magnitude and in the geographical representation of the actual provision of ES. The results also highlight the value of rain gauge and streamflow stations to validate different datasets. A better rain and streamflow gauge network may improve the characterization of the precipitation and streamflow variability in time and space, provide a better validation of precipitation datasets, and may even allow building more local precipitation datasets of higher resolutions. But establishing and maintaining a good pluviometric and streamflow network is expensive and would require several years to have a desirable temporal resolution.

Remote sensing could potentially solve some of these problems in the short term by providing spatial rainfall over large areas (Duan et al., 2012). In the case of CHB, TRMM (algorithm 3B42) appears to be a good option to use as precipitation input. TRMM data are supported both by rain gauge data and by the good fit of the InvEST water yield model as compared to streamflow data. But satellite products such as TRMM have low spatial resolution for modelling experiments targeted at characterizing the provision of ES services at the sub-basin level.

The statistical downscaling of satellite products is an option that could improve the spatial resolution of precipitation gridded datasets. This methodology is based on the relationship between precipitation and other environmental factors, such as topography and vegetation (Jia et al., 2011). Further research is required to evaluate whether this methodology would be useful to ES research in the CHB and other watersheds of Patagonia.

Ecosystem services are increasingly the framework within which conservation and resource management is conducted (Liu et al., 2010; Tallis and Polasky, 2009). While the conceptual appeal of ecosystem services is obvious, the limitations of the most basic physical data such as precipitation and water yield are often overlooked. If ecosystem services are to guide local decisions and management actions, a considerable investment will need to be made in obtaining data on precipitation, which is a primary driver of so many ecological processes.

Acknowledgements

We thank Dr. Peter Kareiva from The Nature Conservancy for his valuable comments all throughout the research. The authors also wish to acknowledge two anonymous reviewers for very helpful suggestions which certainly led to an overall improvement of the paper. This research was funded by the Network for the Conservation of Patagonian River Ecosystems (CONICET and The Nature Conservancy) (Resolution 3213/2).

References

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration, irrigation and drainage 56. FAO, Rome, Italy.
Budyko, M.I., 1974. Climate and Life. Academic, San Diego, CA, USA, pp. 321–330 (translated from Russian by: Miller, D. H.).
Canadell, J., Jackson, R.B., Ehleringer, J.R., Mooney, H.A., Sala, O.E., Schulze, E.-D., 1996. 1996: maximum rooting depth of vegetation types at the global scale. Oecologia 108, 583–595.
Commendatore, M., Esteves, J.L., 2004. Natural and anthropogenic hydrocarbons in sediments from the Chubut River (Patagonia, Argentina). Mar. Pollut. Bull. 48 (9–10), 910–918.
Daly, C., Neilson, R.P., Philips, D.L., 1994. A statistical – topographic model for mapping climatological precipitation over mountainous terrain. J. Appl. Meteorol. 33, 140–158.
Donohue, R.J., Roderick, M.L., McVicar, T.R., 2012. Roots, storms and soil pores: incorporating key ecohydrological processes into Budyko's hydrological model. J. Hydrol. 436–437, 35–50.
Duan, Z., Bastiaanssen, W.G.M., Liu, J.Z., 2012. Monthly and annual validation of TRMM Multisatellite Precipitation Analysis (TMPA) products in the Caspian Sea Region for the period 1999–2003. Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS) 2012, Munich, Germany, July 22–27, pp. 3696–3699.
Ebert, E., Janowiak, J., Kidd, C., 2007. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. Bull. Am. Meteorol. Soc. 88, 47–64. http://dx.doi.org/10.1175/BAMS-88-1-47.
Garreaud, R., Falvey, M., 2009. The coastal winds off western subtropical South America in future climate scenarios. Int. J. Climatol. 20, 543–554.
Garreaud, R., Lopez, P., Minvielle, M., Rojas, M., 2013. Large-scale control on the Patagonian climate.
Getirana, A., Espinoza, J.C., Ronchail, J., RotunnoFilho, O.C., 2011. Assessment of different precipitation datasets and their impacts on the water balance of the Negro River basin. J. Hydrol. 404 (3–4), 304–322. http://dx.doi.org/10.1016/j.jhydrol.2011.04.037.
Giraut, M., Ludueña, S., Lupano, C., Valladares, T.A., 2010. Atlas digital de Cuencas y Regiones Hídricas Superficiales de la República Argentina-Versión 2010. Secretaria de Recursos Hídricos de la República Argentina (CD Rom).
Goyal, R.K., 2004. Sensitivity of evapotranspiration to global warming: a case study of arid zone of Rajasthan (India). Agric. Water Manag. 69, 1–11.
Groisman, P.Y., Legates, D.R., 1994. The accuracy of United States precipitation data. Bull. Am. Meteorol. Soc. 75, 215–227.
Hamel, P., Guswa, A.J., 2015. Uncertainty analysis of a spatially-explicit annual water-balance model: case study of the Cape Fear catchment, NC. Hydrol. Earth Syst. Sci. Discuss. 19, 839–853.
Huffman, G.J., 2012. Algorithm Theoretical Basis Document (ATBD) Version 3.0 for the NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (I-MERG). GPM Project (Greenbelt, MD, 29 pp.) (Available at http://pmm-dev.pps.eosdis.nasa.gov/sites/default/files/document_files/).
Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F., Wolff, D.B., 2007. The TRMM multi-satellite precipitation analysis: quasi-global, multi-year, combined-sensor precipitation estimates at fine scale. J. Hydrometeorol. 8 (1), 38–55.
Insel, N., Poulsen, C., Ehlers, T., 2009. Influence of the Andes Mountains on South American moisture transport, convection, and precipitation. Clim. Dyn. http://dx.doi.org/10.1007/s00382-009-0637-1.
Jia, S.F., Zhu, W.B., Lu, A.F., Yan, T.T., 2011. A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China. Remote Sens. Environ. 115, 3069–3079.
Junzhi, L., A-Xing, Z., Duan, Z., 2012. Evaluation of TRMM 3B42 Precipitation product using rain gauge data in Meichuan Waterhshed, Poyang Lake Basin, China. J. Resour. Ecol. 3 (4), 359–366.
Kareiva, P., Tallis, H., Ricketts, T., Daily, G., Polasky, S. (Eds.), 2011. Natural capital, theory and practice of mapping ecosystem servicesBiology. Oxford University Press.
Legates, D.R., Willmott, C.H., 1990. Mean seasonal and spatial variability in gauge-corrected, global precipitation. Int. J. Climatol. 10, 111–127.
Liang, L., Liu, Q., 2014. Streamflow sensitivity analysis to climate change for a large water-limited basin. Hydrol. Process. 28, 1767–1774.
Liu, S., Costanza, R., Farber, S., Troy, A., 2010. Valuing ecosystem services: theory, practice, and the need for a transdisciplinary synthesis. Ann. N. Y. Acad. Sci. 1185, 54–78. http://dx.doi.org/10.1111/j.1749-6632.2009.05167.x.
Luo, Y., Berbery, E., Mitchell, K., 2005. The operational Eta Model precipitation and surface hydrologic cycle of the Columbia and Colorado Basins. J. Hydrometeorol. 6, 341–370.
Milly, P., 1994. Climate, soil-water storage, and the average annual water-balance. Water Resour. Res. 30 (7), 2143–2156. http://dx.doi.org/10.1029/94WR00586.
Mitchell, T.D., Jones, 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. Int J Climatology vol. 25. Royal Meteorological Society, pp. 693–712. http://dx.doi.org/10.1002/joc.1181.
Nuñez, M., Solman, S., Cabré, M.F., 2008. Regional climate change experiments over southern South America. II: climate change scenarios in the late twenty-first century. Clim. Dyn. 32 (7–8), 1081–1095.
Paturel, J.E., Servat, E., Vassiliadis, A., 1995. Sensitivity of conceptual rainfall-runoff algorithms to errors in input data — case of the GR2M model. J. Hydrol. 168, 111–125.
Plan Director de Recursos Hídricos del Río Chubut, 2013. Informe final, tomo I, documento principal (In Spanish).
Sánchez-Canales, M., López Benito, A., Passuello, A., Terrado, M., Guy, Ziv, Acuña, V., Schuhmacher, M., Elorza, J., 2012. Sensitivity analysis of ecosystem services in a Mediterranean watershed. Sci. Total Environ. 440, 140–153.
Schneider, Udo, Becker, Andreas, Finger, Peter, Meyer-Christoffer, Anja, Rudolf, Bruno, Ziese, Markus, 2011. GPCC Full Data Reanalysis Version 6.0 at 0.5°: Monthly Land-Surface Precipitation From Rain-Gauges Built on GTS-based and Historic Data. http://dx.doi.org/10.5676/DWD_GPCC/FD_M_V6_050.
Sharp, R., Tallis, H.T., Ricketts, T., Guerry, A.D., Wood, S.A., Chaplin-Kramer, R., Nelson, E., Ennaanay, D., Wolny, S., Olwero, N., Vigerstol, K., Pennington, D., Mendoza, G., Aukema, J., Foster, J., Forrest, J., Cameron, D., Arkema, K., Lonsdorf, E., Kennedy, C., Verutes, G., Kim, C.K., Guannel, G., Papenfus, M., Toft, J., Marsik, M., Bernhardt, J., Griffin, R., Glowinski, K., Chaumont, N., Perelman, A., Lacayo, M., Mandle, L., Hamel, P., 2014. InVEST user's guide. The Natural Capital Project (Stanford) (Available at: http://ncp-dev.stanford.edu/~dataportal/invest-releases/ documentation/current_release/, last access: June 2014).
Smith, R., Evans, J., 2007. Orographic precipitation and water vapor fractionation over the Southern Andes. J. Hydrometeorol. 8, 3–19.
Tallis, H., Polasky, S., 2009. Mapping and valuing ecosystem services as an approach for conservation and natural-resource management. The year in ecology and conservation biology. Ann. N. Y. Acad. Sci. 1162, 265–283. http://dx.doi.org/10.1111/j.1749-6632.2009.04152.x.
Xu, X., Liu, W., Scanlon, B.R., Zhang, L., Pan, M., 2013. Local and global factors controlling water-energy balances within the Budyko framework. Geophys. Res. Lett. 40, 6123–6129.
Xu, C., Tunemar, L., Chen, Y., Singh, V.P., 2006. Evaluation of seasonal and spatial variations of lumped water balance model sensitivity to precipitation data errors. J. Hydrol. 324 (1–4), 80–93.
Yilmaz, K., Hogue, T., Hsu, K., Sorooshian, S., Gupta, H., Wagener, T., 2005. Intercomparison of rain gauge, radar and satellite-based precipitation estimates with emphasis on hydrologic forecasting. J. Hydrometeorol. 6 (4), 497–517.